
Conference or Workshop Item:
Oxnard, Stephen and Murphy, Damian Thomas orcid.org/0000-0002-6676-9459 (2014) Achieving Realistic Auralisations Using an Efficient Hybrid 2D Multi-plane FDTD Acoustic Model. In: EAA Joint Symposium on Auralization and Ambisonics, 03-05 Apr 2014.

https://doi.org/10.14279/depositonce-21

ACHIEVING REALISTIC AURALISATIONS USING AN EFFICIENT HYBRID 2D MULTI-PLANE FDTD ACOUSTIC MODEL

Stephen Oxnard, *

University of York
Audio Lab, Department of Electronics
York, UK
so523@york.ac.uk

Damian Murphy,

University of York
Audio Lab, Department of Electronics
York, UK
damian.murphy@york.ac.uk

ABSTRACT

This research examines the validity of utilising a 2D multiplane FDTD acoustic model to simulate low frequency sound propagation as part of a hybrid room impulse response (RIR) synthesis system. Analytic results, pertaining to the comparison of simulated low frequency multiplane RIRs with both practical RIR measurements and 3D FDTD simulated RIRs, demonstrate that a good level of accuracy is attained through use of this hybrid modelling paradigm. This claim is further supported, in part, by comparative subjective test results. Furthermore, 2D multiplane simulations are shown to be far more efficient than full 3D FDTD modelling procedures as they achieve a $\sim$98% reduction in computation time.

1. INTRODUCTION

Traditionally, geometric acoustic modelling approaches have been harnessed to virtually simulate sonic environments for the purposes of room acoustics prediction and analysis. These approaches, which include ray-tracing, image source method (ISM) and acoustic radiosity (AR), while computationally efficient, produce an inaccurate representation of sound propagation at low frequencies. This arises due to underlying assumptions, common to the implementation of all geometric modelling techniques, which do not facilitate the preservation of wave phenomena such as interference effects, diffraction and standing waves. Conversely, modern wave-based approaches, such as the Finite Difference Time Domain (FDTD) and Digital Waveguide Mesh (DWM) paradigms, allow for direct numerical solution of the wave equation. As such, wave-based acoustic models inherently emulate the behaviour of low frequency sound propagation to a far greater level of accuracy than their geometric predecessors. However, the implementation of these numerical acoustic models is hindered greatly by their reliance on extensive computational resources and lengthy compute times.

Despite the computational challenges posed by numerical approaches, which are particularly apparent for 3D acoustic simulations, implementations of 3D FDTD models may be processed in real-time under certain conditions. A previous study [1] demonstrates that it is possible to render 3D soundfields by means of the FDTD paradigm at interactive sampling rates through utilisation of Graphics Processing Units (GPUs). However, this outcome is realisable only when restricting the size of the modelled spatial domain and/or the simulated frequency bandwidth. An alternative, efficient wave-based modelling method devised by Raghuvanshi et al. [2, 3] applies Adaptive Rectangular Decomposition (ARD) to segment the spatial domain into cuboid sections for which the analytical solution of the wave equation is known. As such, the entire modelled soundfield may be resolved temporally through a series of weighted Cosine basis functions. In [3], the ARD system is shown to achieve significant savings for band-limited acoustic modelling in terms of memory consumption and requires up to 18x less computation compared to FDTD schemes, noting that a GPU implementation was not used. Further reductions in computational expense are attained when reducing the spatial dimensionality of the model. This was discussed in relation to room acoustic modelling by Kelloniemi et al. in [4] where multiple 2D DWMs were used to emulate acoustic simulations of a geometrically simplistic 3D soundfield. In this work, the authors predicted reductions in memory consumption and processing cost of 97% and 99% respectively against a 3D DWM model while preserving important spectral features present in the soundfield.

All numerical acoustic modelling implementations are limited in their use in terms of valid bandwidth due to inherent dispersion error which becomes evident with increasing frequency [5]. For this reason, full bandwidth wave-based models rely on greatly over-sampled spatial domains leading to large increases in computational cost. In attempt to alleviate the trade-off between accuracy and efficiency in full audio bandwidth room acoustic simulations, several examples of hybrid modelling systems have been developed. These methods seek to render room impulse responses (RIRs) through complementary assimilation of two or more virtual modelling paradigms. An example of such a system, documented in [6] combined use of a wave-based 3D FDTD scheme with optimised ISM and AR models for hybrid RIR generation. This modelling approach limited the application of FDTD simulations to low frequencies and, hence, reduced the required computation load and run-times comparative to full audio bandwidth FDTD schemes. Results obtained were then amalgamated with high frequency RIRs generated by the ISM and AR models to render a spectrally complete hybrid RIR. A more recent study [7], largely influenced by the research of [4], sought to validate the use of multiple 2D cross-sectional FDTD schemes as a means of representing low frequency sound propagation throughout a 3D enclosure, again as part of a hybrid modelling approach. Results presented in this work demonstrated that the wave-based 2D multiplane approach achieved a reasonable approximation to low frequency RIRs simulated by a full 3D FDTD scheme, in a simplistic modelling scenario, while reducing simulation run-times by 99.15%.

The research documented in this paper seeks to further examine the validity of utilising 2D multiplane FDTD schemes as an

* Funded by a U. of York EPSRC Doctoral Training Grant.

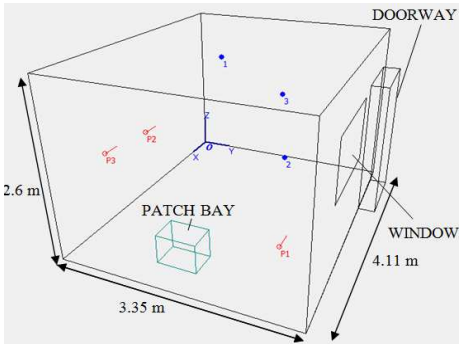

Figure 1: 3D depiction of the Live Room geometry (as viewed in ODEON) highlighting significant dimensions and features and the orientation with the x,y and z axes. Source and receiver placements are represented by 'P1-P3' and '1-3' respectively.

efficient means of low frequency acoustic modelling. This task was undertaken by utilising the multiplane approach, in conjunction with ray-based geometric techniques, to render hybrid RIRs of an existing space. Resulting low frequency RIR spectra were then compared against those of practical RIR measurements and those generated by a 3D FDTD model of the enclosure in order to analyse the level of agreement between both virtual models and reality. Finally, preliminary subjective tests were conducted to compare auralisations rendered by means of 2D multiplane and 3D FDTD hybrid acoustic modelling in a perceptual sense. It is proposed that 2D multiplane FDTD modelling will offer a highly efficient means of augmenting geometric approaches to rendering large-scale acoustic scenes for virtual reality and acoustic prediction applications where perceptually valid, real-time or interactive design workflows are required.

2. THE VIRTUAL ACOUSTIC MODELS

The acoustic models created for the purposes of this study were virtual representations of the recording studio live room situated in the Audio Lab at the University of York. An overview of the room geometry is depicted in Figure 1. Practical RIR measurements were obtained from this space prior to the modelling procedure in order to gain real results for comparison. For the purposes of RIR measurement, an omni-directional sound source was approximated by rotating a Genelec 8130A loudspeaker around the azimuth at increments of 90°and capturing IRs for each orientation using a ST450 Soundfield microphone. Mono RIRs were rendered by summing the W-channel of the captured B-Format responses obtained for each loudspeaker orientation. In this way, three impulse responses were collected from the live room using the sound source and receiver placements detailed in Table 1. In total, three acoustic models, described here, were constructed to simulate the acoustic of this space for each case of source and receiver location, yielding three RIRs per model.

2.1. Geometric Model

ODEON 10.1 Auditorium [8] acoustic prediction software was utilised to develop the geometric acoustic model of the live room and render the mid-high frequency RIRs. This software package is an industry standard acoustic modelling program which combines

CASE	Source (x,y,z) (m)	Receiver (x,y,z) (m)
1	(3.61, 2.85, 0.68)	(0.60, 0.69, 2.00)
2	(2.91, 0.65, 1.49)	(2.91, 2.65, 1.49)
3	(3.61, 0.50, 1.45)	(0.70, 1.85, 1.50)

Table 1: Overview of source and receiver placements in each RIR measurment case.

ray-tracing and ISM to render RIRs for virtual environments developed by the user.

The geometries of the live room were compiled from architectural diagrams and input to ODEON via a .par (parameter) file. Sound source and receiver locations were defined within the model with reference to the locations utilised for practical measurements. Reference receivers were also defined at a distance of 1m from each sound source location for the purposes of the RIR calibration process described in section 3. A total of 7 different material types were applied to the surfaces incorporated in the model: Floor (carpet); Ceiling (fiber board); Walls (plaster board and cavity); Window (double glazed); Door (solid wood); Patch Bay outer shell (medium density wood); Patch Bay front panel (metal). The closest approximate material properties available in the ODEON material library were applied as appropriate. To generate results of suitable accuracy, 50000 rays were used to render each impulse response. Mono omni-directional RIRs were obtained by extracting the W-channel of B-Format impulse responses in each case in order to maintain consistency both with practical measurements and with the capture method implemented in the FDTD schemes.

2.2. 3D FDTD Model

The derivation of the 3D FDTD scheme utilised in this work begins with the 2^{nd} order homogeneous wave equation:

$$\frac{\partial^2 p(\overrightarrow{n},t)}{\partial t^2} = c^2 \nabla^2 p(\overrightarrow{n},t) \tag{1}$$

where p is a measure of acoustic pressure at time t at a location given by the 3D Cartesian positional vector $\overrightarrow{n}$, c is wave speed in ms^{-1} and ∇^2 is the 3D Laplacian operator. Spatio-temporal discretisation of (1) through use of centered finite difference approximations yields the 'Standard Rectilinear' (SRL) update equation, see e.g. [9], implemented in this study. To ensure the FDTD scheme operates with numerical stability a lower limit is imposed on the magnitude of the spatial sampling interval. For 3D FDTD schemes this lower limit h_{3D}, calculated by means of Von Neumann analysis [10], is given as $h_{3D} \geq ck\sqrt{3}$ where k is the discrete time step (s) defined as the reciprocal of the temporal sampling frequency Fs. The spatial discretisation of (1) inherently gives rise to dispersion error, due to anisotropic wave propagation, which is most apparent at high frequencies. In order to reduce the impact of dispersion effects, h_{3D} was set equal to the lower limit giving a usable simulation bandwidth of 0.196Fs [9]. Frequency independent locally reactive surface (LRS) boundary conditions [9] were utilised to terminate the spatial domain allowing for appropriate reflection coefficients to be applied to each bounding surface. Suitable reflection coefficients for each modelled boundary surface were derived by averaging the low frequency absorption coefficients, provided in ODEON, for the corresponding surface material type applied in the geometric model. Table 2 provides the reflection coefficients applied to each surface in the 3D FDTD

Surface	R_{3D}	$R_{2D,MP}$
Floor	0.8860	0.7982
Walls	0.9592	0.9061
Ceiling	0.8944	0.8541
Window	0.9747	0.9442
Door	0.9539	0.9033
Patch Bay Shell	0.9487	0.9002
Patch Bay Front	0.7416	0.6742

Table 2: Overview of reflection coefficients applied to different surface types in the 3D (R_{3D}) and 2D multiplane ($R_{2D,MP}$) FDTD models .

model.

The 3D FDTD model may be envisaged as a rectilinear lattice of pressure sampling nodes occupying the volume defined by the dimensions of the live room with an internodal distance of 0.0135m corresponding to Fs = 44.1kHz. The positioning of boundary surfaces and source and receiver placements within the model was calculated by rounding the actual position to the nearest sampling instance with a maximum deviation of $((3h_{3D})^2/4)^{0.5} < 0.012$m. A Dirac Delta sound excitation signal was applied for each RIR case by initialising the source node to a unity value. Impulse responses were then captured by recording the response to this excitation at the receiver node. Approximately 17 000 000 sampling nodes were required to render the 3D model and average computation times were in the region of 3800s for 1s RIRs at the audio sampling rate.

2.3. 2D Multiplane FDTD Model

The 2D Multiplane FDTD model was designed to approximate the 3D live room geometry through use of three intersecting 2D SRL FDTD schemes derived in an analogous manner to the 3D scheme with $\overrightarrow{n}$ and ∇^2 of (1) reduced to 2 dimensions. Again, a lower limit was imposed on the spatial sampling interval in the interest of maintaining numerical stability, which for the 2D FDTD case is: $h_{2D} = ck\sqrt{2}$. This gives an inter-nodal distance of 0.011m for Fs = 44.1kHz yielding a maximum spatial positioning error of source, receiver and bounding surface placement of $((h_{2D})^2/2)^{0.5} < 0.008$m comparative to practical dimensions. Locally reacting surface boundary conditions were also utilised in the multiplane model, however the reflection coefficients applied to each bounding surface (see Table 2) were subject to calibration as discussed in section 3.2.

Each 2D FDTD scheme represented a cross-section of the space orientated in the x-y, x-z and y-z planes with a common point of intersection defined by the position of the receiver in each measurement case. Sound excitation positions were defined by taking the sound source locations and projecting them perpendicularly onto each plane. In this way, the excitation nodes defined in the x-y, x-z, and y-z orientated cross-sectional schemes shared the (x,y), (x,z) and (y,z) coordinates of the measurement source locations respectively. As with the 3D case, excitation of each plane was achieved by initialising each source node to a unity pressure value. The individual responses captured from the cross-sectional planes were aligned in time, synchronising the direct sound component of each 2D RIR, and then summed to obtain the complete multiplane RIR.

The multiplane models created for each RIR measurement case consisted of approximately 300 000 nodes, which equates to less than 2% of the number of nodes required for the 3D FDTD model. Hence, a very large computational saving is gained through use of the multiplane model. This is further reflected in comparative computation times. For the generation of a 1s RIR at audio rate, the multiplane models required a processing time of approximately 60s, therefore achieving a run-time reduction of $\sim$98% compared against the 3D model.

3. VIRTUAL RIR CALIBRATION AND HYBRIDISATION

3.1. 3D FDTD RIR Calibration

Once the 2D and 3D FDTD RIRs have been obtained they are processed such that they can be combined with the mid-high frequency geometric (GA-) RIRs generated in ODEON. In the case of 3D FDTD RIRs, this task was undertaken in accordance with the RIR matching procedure documented in [6]. The initial step involves calibrating each GA-RIR by equalising the total acoustic energy recorded at the corresponding reference receiver, positioned at a distance of 1m from each sound source in the geometric model, to a unity value. To this end, the total energy E_T of the GA reference impulse response captured at each reference receiver was calculated using,

$$E_T = \sum_{n=1}^{N} p^2[n] \tag{2}$$

where $p[n]$ is the pressure value recorded at temporal sampling instant n and N is the length of the impulse response in samples. To proceed, a constant K, which may be applied to reduce the total IR energy to unity, is calculated using the following relation:

$$K = \sqrt{\frac{1}{E_T}} \tag{3}$$

As such, three values of K were calculated and applied to the corresponding GA-RIRs through multiplication, thus appropriately calibrating the geometric model for each response case. The resulting GA-RIR signals were then high pass filtered to remove all spectral components below a cut-off frequency of 2kHz. This cut-off frequency was selected to ensure that the results obtained from FDTD modelling, applied to create the hybrid RIR, consisted of measurements well within the usable bandwidth of the numerical scheme computed.

In order to avoid reducing the effectiveness of the energy matching procedure between geometric and FDTD results, it was first necessary to process the 3D FDTD RIRs and remove D.C. components arising due to the nature of the excitation function and erroneous pressure recordings occurring at mid-high frequencies due to dispersion effects. As such, each 3D FDTD RIR was passed through a 2^{nd} order D.C. blocking filter and a low pass filter with a cut-off of 2kHz. Having done so, it was then possible to calibrate the 3D FDTD RIRs with the geometric results by multiplying each response with an energy matching parameter n, defined in [6] as:

$$n = (5.437 \times (Fs \times 10^{-3})) - 3.6347 \tag{4}$$

The spectrally complete hybrid RIRs were then created by summing the corresponding fully calibrated geometric and 3D FDTD impulse responses.

3.2. 2D Multiplane FDTD RIR Calibration

Due to dissimilarities of the laws governing energy decay in 2D and 3D FDTD schemes, as documented in [11], it was expected that the multiplane model would exhibit comparatively longer decay times. For this reason, calibration of reflection coefficients applied to the boundary of each 2D plane was necessary in order to obtain RIR reverberation times (RT_{60}) consistent with those generated in the 3D model. To this end, a series of simple cubic 2 x 2 x 2 m 3D FDTD lattices were constructed to obtain the $RT_{60,3D}$ of each cube with constant reflection coefficients, equal to those utilized in the 3D live room model, applied to all surfaces. The recorded $RT_{60,3D}$ values were then inserted to the Norris-Eyring for 2D RT_{60} rearranged to make the subject the boundary absorption value α:

$$\alpha = 1 - \exp\left(\frac{-\pi Sln(10^6)}{cLRT_{60,3D}}\right) \tag{5}$$

where S and L were set to the surface area and side length of the 2 x 2 m cross-section of the 3D cubic schemes. Using (5) it was possible to calculate appropriate boundary absorption values, and hence reflection coefficients, for the 2D multiplane model that corresponded to those applied in both the geometric and 3D FDTD models.

The RIRs simulated by the multiplane model were filtered in a similar manner to the 3D FDTD RIRs in order to remove D.C. components and frequencies above 2kHz. In addition, a further filtering stage was required to remove the effects of afterglow, as per the procedure documented in [12]. This was appropriate as the afterglow phenomenon in 2D numerical schemes acts to erroneously skew the magnitude of low frequency spectral components.

The final calibration stage involved matching the total energy present in the filtered multiplane RIRs to that of the corresponding calibrated 3D FDTD RIRs and, by extension, the geometric RIRs. This was carried out by applying (2) to the multiplane RIRs and multiplying each RIR by a matching constant K_n defined as follows:

$$K_n = \sqrt{\frac{E_{T,3Dn}}{E_{T,2Dn}}} \tag{6}$$

where $E_{T,2Dn}$ and $E_{T,3Dn}$ are the total energies of the 3D and 2D multiplane RIRs, respectively, for each measurement case $n = 1, 2, 3$. Having been calibrated by K_n the multiplane RIRs were then summed with geometric results to produce the complete 2D multiplane FDTD/geometric hybrid impulse responses. Currently, the correct calibration of the 2D multiplane RIRs relies on the generation of corresponding 3D RIRs for use as a reference. It is intended that this reliance will be removed in future work through development of an analytical means of matching the energy levels present in GA-RIRs and multiplane RIRs closely following the work presented in [6] and [13].

4. OBJECTIVE RESULTS

The results documented in this section are derived from the measured RIRs and those simulated by the 2D and 3D FDTD hybrid approaches. For the purposes of this study, analysis of RIRs is constrained to low frequency spectra in order to evaluate the results generated by the 2D multiplane and 3D modelling methods and then to compare both sets of results with real acoustic data.

4.1. Low Frequency Analysis

Three impulse responses, representative of the three measurement configurations applied in practice (see Table 1), were collected from each hybrid acoustic model. Figure 2 depicts a graphical comparison of the low frequency magnitude responses of the measured RIRs and the 2D multiplane and 3D FDTD hybrid RIRs in each measured case. For all cases, it is apparent that the magnitude of the low frequency response obtained in practice (denoted 'real') is consistently greater than that of either the virtual RIRs. This discrepancy is simply due to a difference in sound excitation strength applied in the real and virtual environments and was rendered negligible by means of a normalisation procedure prior to construction of material used during subjective testing later described. The magnitude responses generated by both hybrid models appear to show good agreement in overall energy levels which exposes the success of the energy matching procedure previously discussed.

The spectra presented for measurement cases 1 and 2 depict good agreement between low frequency components of measured RIRs and RIRs produced by both models below 150Hz. In this frequency range, comparable alignment of resonant peaks is observed, suggesting that strong axial modes are well represented by both acoustic models. Referring to case 1 in the range of 160-200Hz, it is apparent that the multiplane model does not accurately recreate the modal aspects present in both measured and 3D modelled spectra. This is potentially due to the inherent inability of the multiplane model to capture prominent oblique modes occurring in this frequency range. However, a similar disparity is not observed in case 2, where the multiplane model achieves a far better representation of the notch (190Hz) occurring in the measured response than the full 3D model. The cause of this result is yet to be investigated. Beyond 200Hz, in both cases 1 and 2, a reasonable correlation exists between the multiplane, 3D and measured RIR spectra in terms of spectral component positioning. Measured low frequency behaviour in case 3 is shown to be better represented by the synthesised 3D RIR than that of the multiplane models, however, agreement between all spectra is notable below 100Hz. In this case the 2D multiplane response exhibits a lack of clarity in terms of defined resonant components above 200Hz comparative to measured and 3D modelled results. In summary, the overall representation of low frequency characteristics possessed by synthesised 3D RIRs and measured RIRs attained by the multiplane model is very encouraging considering the achieved reduction in run-time compared to full 3D FDTD modelling (see section 2.3).

4.2. Global Reverberation and Early Decay Times

Reverberation time T_{30} and early decay times (EDT) were derived from the RIRs captured in the live room and the virtual models for each measurement scenario in accordance with ISO documentation [14]. As such, it was possible to calculate global values for both parameters by averaging the values returned for each measurement case applied in both models and in practice. Table 3 provides a review of the recorded parameter values in the 500Hz frequency octave. The additional 'Just Noticeable Difference' (JND) ranges were calculated by applying 5% JND [15] for EDT and 30% JND for T_{30} [6]. These percentage values refer to the maximum deviation from the true value beyond which the difference in decay times become perceptible. Hence, the JND measures provided offer insight to the subjective tolerance range for EDT and reverberation time with reference to the value calculated for each environment.

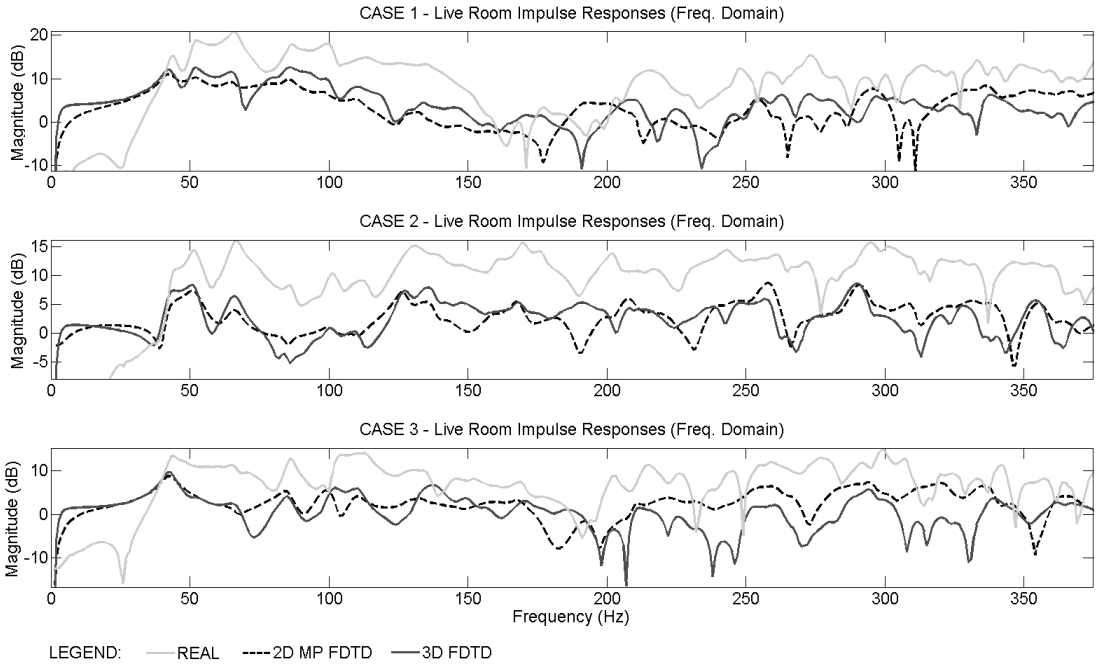

Figure 2: Comparison of magnitude spectra of real (solid grey), 2D Multiplane (black dashed) and 3D FDTD (solid black) RIRs at low frequencies for all three cases of source and receiver combinations.

	$T30_{500Hz}$ (s)	$T30$ JND Range (s)	EDT_{500Hz} (s)	EDT JND Range (s)
2D MP FDTD	0.38	0.27 - 0.49	0.26	0.25 - 0.27
3D FDTD	0.36	0.25 - 0.47	0.41	0.39 - 0.43
Real	0.44	0.31 - 0.57	0.42	0.40 - 0.44

Table 3: Global $T30_{500Hz}$ and EDT_{500Hz} values derived from RIRs measured in practice ('Real') and those rendered by the 2D multiplane and 3D FDTD hybrid acoustic models ('2D MP FDTD' and '3D FDTD' respectively).

As shown in Table 3, good agreement is exhibited between the 2D Multiplane and 3D FDTD model reverberation times with a minimal discrepancy of 0.2s. This demonstrates the effectiveness of the reflection coefficient matching procedure applied to the multiplane model (see section 3.2). A similar result is not observed for EDT where the difference between 2D and 3D models is much greater than JND tolerance values. This outcome is unexpected as energy levels, due to an impulse excitation, in 2D FDTD schemes should take longer to decay than those in 3D. Hence, it may be hypothesised that the values returned for the particular octave band under examination may not be representative of the overall EDT characteristics of the multiplane RIRs, however this claim requires further investigation. In addition, EDT is dependent on the distribution and amplitude of early reflections. In the case of the 2D multiplane model, low frequency reflection paths are represented only in planar cross-sectional areas as opposed to the volume of the modelled space in its entirety. As such, the temporal density of early reflections present in the multiplane FDTD RIRs is expected to be less than that of the 3D FDTD RIRs. This issue, which might also influence the EDT results presented in Table 3, remains to be examined in future work. Through comparison of the live room T_{30} with those of the models, it is apparent that reverberation times simulated in the virtual models lie within the JND range of the live room T_{30}, and hence the audible discrepencies should be negligible. With reference to the EDT measured in the live room, the 3D FDTD model is shown to produce the closest approximation of initial sound decay characteristics.

5. SUBJECTIVE TESTING

A simple preliminary listening test, described in the following, was constructed in order to support or disprove the following hypothesis:

"Auralisations generated by means of 3D FDTD/geometric and 2D Multiplane FDTD/geometric hybrid modelling will exhibit agreeable levels of similarity to auralisations rendered through use of measured RIRs in terms of perceived frequency response and reverberation."

5.1. Listening Test Material and Procedure

In review, a total of 9 RIRs were captured during the course of this study: 3 from measurements of the live room, and 3 from each of the hybrid models, representing 3 source and receiver configurations applied in practice. These RIRs were utilised to create aural-

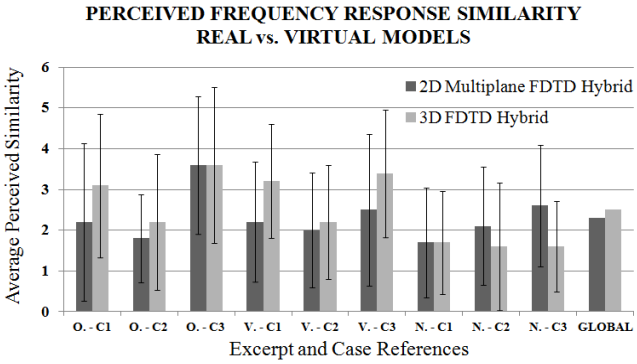

Figure 3: Average perceived similarity of frequency response between real and virtual auralisation for all excerpt and case combinations. Global averages are also shown.

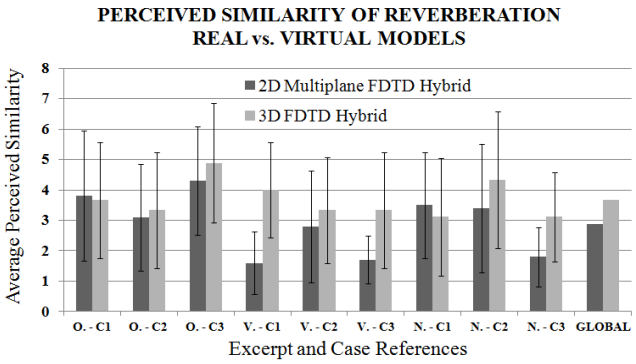

Figure 4: Average perceived similarity of reverberation between real and virtual auralisation for all excerpt and case combinations.

isations of the real and modelled spaces by convolving each with 3 anechoic audio excerpts: an orchestral recording (25s), an all-male vocal quartet (7s) and a pink noise burst (2s). This resulted in an audio test set of 27 auralisations, 9 of which were of the live room corresponding to the performance and capture of each anechoic excerpt as per all three cases of source and receiver placements. The remaining 18 auralisations were the virtual representations of the 9 auralisation scenarios as rendered by each hybrid acoustic model.

This auralisation set was arranged into 18 pairs such that each pair consisted of an auralisation rendered from live room measurements ('real') and the corresponding auralisation produced by means of one of the two hybrid acoustic models ('virtual'). To clarify, 9 test pairs compared the real auralisations with those rendered from 3D FDTD/geometric modelling and the remaining 9 compared real auralisations with those rendered from 2D Multiplane FDTD/geometric modelling.

During the test, subjects were presented with the auralisations in 'AB' pairs where, in each instance of the 18 pairs, 'A' was randomly selected as a real auralisation while 'B' was a virtual auralisation, or vice versa. Subjects were asked to listen to and compare the level of similarity between auralisation 'A' and 'B' with reference to frequency response and reverberation. Results were recorded by enabling subjects to rate the similarity of each perceptual parameter on a 7-point 'scale of similarity' on which a returned value of '0' corresponded to 'highly dissimilar' and '6' corresponded to 'highly similar'. This process was repeated for all 18 pairs of auralisations.

Due to the composition of the measured and rendered RIRs, the auralisations were rendered in Mono audio format and presented to subjects over headphones with both left and right channels producing the same signal. Volume levels remained consistent across all tests and the ordering of presented auralisation pairs was randomised for each subject.

5.2. Listening Test Results

The results provided by 10 test participants were combined and averaged to find the mean rating of similarity in terms of frequency response and reverberation for each model, case and excerpt combination. In doing so it was possible to compare the level to which each hybrid model compared with the auralisations of the live room and, hence, discern the effectiveness of the 2D multiplane FDTD hybrid model comparative to that of the 3D FDTD hybrid system, as per the test hypothesis.

Figure 3 displays the average perceived similarity of frequency response between both virtually modelled and real auralisations. For brevity, excerpts are denoted O., V., and N. for orchestra, voice and noise respectively. Likewise, cases 1-3 are denoted C1, C2 and C3. The error bars shown describe the standard deviation of recorded results from all tests with respect to the mean similarity values. As such, it may be observed that the perception, or rating, of similarity varied significantly between subjects for each auralisation comparison. However, from the mean values it is clear that, in terms of frequency response, the 3D and multiplane hybrid models performed comparably for O. - C3 and N. - C1. In contrast, the 3D hybrid model auralisations were perceived as most similar to the live room auralisations for the vocal excerpt in all measurement cases. This is most likely due to the content of the vocal excerpt which consisted of a 4-part male vocal ensemble and, therefore, was comprised of mainly low-mid frequencies. Referring back to Figure 2, it is shown that the multiplane responses at low frequencies best match those of the 3D model responses for measurement case 2. This is reflected in the results for the vocal auralisations suggesting that for auralisations, comprised of mostly low frequencies, perceptual similarities in frequency response may be more dependent on the placement and definition of resonant spectral components than initially thought. This interesting outcome is to be examined in future work in order to verify whether or not this is the case. Subjective comparison between virtual model auralisations and real auralisations rendered using the pink noise excerpt suggests that the multiplane model exhibits a higher level of realism than the 3D model (N. - C2 and C3) in terms of perceived frequency response. This outcome, which is converse to that demonstrated for the vocal auralisations, demonstrates that 2D multiplane modelling may be more appropriate than 3D FDTD modelling for particular auralisation purposes. Lastly, global test results noted in figure 3 appear to support the subjective test hypothesis. These global values were calculated by taking the mean value of all average similarities shown in figure 3 for each virtual model. In doing so, the influences of the audio excerpts and RIR measurement cases on the test results were bypassed to gain an overall measure of similarity. The discrepancy between these values for the 2D multiplane and 3D FDTD hybrid models is encouragingly small being in the region of 0.21.

In figure 4, the average perceived similarity between the live room auralisations and those produced using the two hybrid acoustic models is displayed in terms of reverberation. Considering first the comparison of real and virtual auralisations produced using the vocal excerpt, it may be observed that the 3D model auralisations consistently outperform those of the multiplane model with respect to perceived reverberance. As previously noted, this particular excerpt may expose the inaccuracies of the multiplane approach and, hence, the disparity between measured EDTs noted in Table 3 may be impacting these results. For the orchestral auralisations, discrepency between the reveberance of each model is shown to be small. The similarity ratings for each model are variable in the case of the pink noise excerpt auralisations, however global results (calculated analogously to those in figure 3) suggest a reasonably small comparative difference between modelling approaches in terms of perceived reverberance and, hence, provide partial support of the test hypothesis.

It is noted that the overall perceived similarity between model auralisations and auralisations rendered from practical RIR measurements is relatively low, as shown by global results for both frequency response and reverberation. This is expected considering the assumptions on which both hybrid modelling paradigms are based and the fact that significant acoustic phenomena, such as resonances in bounding surfaces, are not as yet possible to model. Moreover, particular aspects of the hybrid acoustic models could be refined in order to better represent a real acoustic environment. Such refinements would see the inclusion of approximated sound energy attenuation due to viscosity of air, frequency dependent boundary absorption characteristics at low frequencies and more accurate representations of sound source/receiver frequency response and directivity characteristics.

6. CONCLUSION

This paper provides an overview, and assesses the performance, of a recently devised 2D multiplane FDTD hybrid modelling paradigm applied to a realistic acoustic modelling scenario. Objective results obtained from RIR measurements demonstrate that this efficient multiplane approach possesses the potential to model realistic low frequency sound fields to a level of accuracy similar to that of 3D FDTD schemes while reducing run-times by approximately 98%. Results generated by conducting subjective listening tests support the claim that 2D multiplane and 3D FDTD hybrid models produce comparable levels of realism in rendered auralisations when compared against auralisations generated from practical measurements. With reference to the findings of this study, it is proposed that future work will aim to investigate and contrast the representation of early reflections in 2D and 3D FDTD acoustic modelling to better match RIRs resulting in each case. Additionally, the applicability of the mutliplane technique for acoustic simulation will continue to be assessed through a series of further subjective tests.

7. REFERENCES

[1] L. Savioja, "Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics," in *Proc. of the 13th Int. Conf. on Digital Audio Effects (DAFx-10)*, Graz, Austria, 2010.

[2] N. Raghuvanshi, B. Lloyd, N. K. Govindaraju, and M. C. Lin, "Efficient numerical acoustic simulation on graphics processors using adaptive rectangular decomposition," in *Proc. of the EAA Symp. on Auralization*, Espoo, Finland, 2009.

[3] N. Raghuvanshi, R. Narain, and M. C. Lin, "Efficient and accurate sound propagation using adaptive rectangular decomposition," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 5, pp. 789–801, 2009.

[4] A. Kelloniemi, V. Valimaki, and L. Savioja, "Simulation of room acoustics using 2-D digital waveguide meshes," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, 2006, pp. 163–168.

[5] L. Savioja and V. Valimaki, "Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 2, pp. 184–194, 2000.

[6] A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja, "Room Impulse Response Synthesis and Validation Using A Hybrid Acoustic Model," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 9, pp. 1940–1952, 2013.

[7] S. Oxnard and D. Murphy, "Room impulse response synthesis based on a 2D multi-plane FDTD hybrid acoustic model," in *Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, 2013.

[8] ODEON Website. (2014) Last Accessed 15th Feb. [Online]. Available: http://www.odeon.dk

[9] K. Kowalczyk and M. van Walstijn, "Room Acoustics Simulation Using 3-D Compact Explicit Schemes," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 1, pp. 34–46, 2011.

[10] S. A. V. Duyne and J. O. Smith III, "Physical Modeling with the 2-D Digital Waveguide Mesh," in *Proc. of the Int. Conf. on Computer Music*, San Francisco, CA, 1993, pp. 40–47.

[11] J. Wells, D. Murphy, and M. Beeson, "Temporal matching of 2D and 3D wave-based acoustic modeling for efficient and realistic simulation of rooms," in *Audio Eng. Soc. (AES) Conv. 126*, Munich, Germany, 2009.

[12] J. Escolano, C. Spa, A. Garriga, and T. Mateous, "Removal of afterglow effects in 2-D discrete-time room acoustics," *J. Applied Acoustics*, vol. 74, no. 6, pp. 818–822, 2013.

[13] S. Siltanen, A. Southern, and L. Savioja, "Finite-difference time domain method source calibration for hybrid acoustics modeling," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, 2013, pp. 166–170.

[14] European Standard Documentation, "Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces," *ISO 3382-1:2009*, 2009.

[15] N. Prodi and S. Velecka, "The evaluation of binaural playback systems for virtual sound fields," *J. Applied Acoustics*, vol. 64, no. 2, pp. 147–161, 2003.